Attorney's Docket No.: Enz-60(CIP)

APPLICATION

FOR

UNITED STATES LETTERS PATENT

To Whom It May Concern:

Be it known that

**Elazar Rabbani, Jannis G. Stavrianopoulos, James J. Donegan, and Jack Coleman**

have invented certain new and useful improvements in

# NOVEL COMPOSITIONS AND METHODS FOR CONTROLLING THE EXTENDABILITY OF VARIOUS COMPONENTS USED IN COPYING OR AMPLIFICATION STEPS

Of which the following is a full, clear and exact description.

---

**EXPRESS MAIL CERTIFICATE**

"Express Mail" Label No. EV151099758US
Deposit Date October 24, 2003
I hereby certify that this paper and the attachments herein are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington DC 20231.

Natalie Bogdanos — October 24, 2003
Natalie Bogdanos — Date
Reg. No. 51,480

# NOVEL COMPOSITIONS AND METHODS FOR CONTROLLING THE EXTENDABILITY OF VARIOUS COMPONENTS USED IN COPYING OR AMPLIFICATION STEPS

## REFERENCE TO RELATED PATENT APPLICATIONS

This application is a continuation-in-part of U.S. Application No. 09/896,897, filed on June 30, 2001, entitled "Novel Compositions and Processes for Analyte Detection, Quantification and Amplification." The content of the aforementioned application is hereby incorporated by reference, in its entirety.

## FIELD OF THE INVENTION

This invention relates to novel compositions and methods for controlling the extendability of various components used in copying or amplification steps to provide an enhancement of the specificity of the nucleic acids produced by these processes as well as potentially increasing the amount of such products.

All patents, patent applications, patent publications, scientific articles and the like, cited or identified in this application are hereby incorporated by reference in their entirety in order to describe more fully the state of the art to which the present invention pertains.

## SUMMARY OF THE INVENTION

In the present invention, methods and compositions are disclosed that allow control of the extendibility of nucleic acids. By these means, the copying or amplification of target nucleic nucleic acids can be carried out under conditions where synthesis (derived from target nucleic acid templates) is retained while potentially deleterious side reactions caused by nucleic acids acting inappropriately as either primers or as templates are avoided.

In some aspects of the present invention, methods are disclosed for treatment of nucleic acids in a sample that allow the nucleic acids to be used as templates but prevent or inhibit their functioning as primers themselves. Methods are given for the treatment of both RNA and DNA targets. The target nucleic acids may be isolated from biological samples or they may be copies of nucleic acids isolated from biological samples.

In another aspect of the present invention, novel primers are disclosed where the nature of nucleotides at the 3' end of the primers are selected for an ability to carry out template dependent nucleic acid synthesis while at the same time for an inability to participate in a template independent extension reaction if they have not been previously extended through the use of a target template.

For instance, when primers are used to create copies of target nucleic acids, if a further reaction is carried out by terminal transferase to add a homopolymeric sequence, extended primers comprising copies of target nucleic acid sequences will still be substrates for terminal transferase addition, while excess unextended primers that may be present in the reaction mixture will be refractory to terminal transferase addition. By these means, undesirable products from unextended primers participating in further biochemical reactions, such as serving as templates or providing transcription complexes, are avoided.

As such, this process should allow more efficient synthesis of desirable sequences.

In another aspect of the present invention, methods and compositions are disclosed that allows a rapid and efficient addition of desirable nucleic acid sequences to the ends of single stranded nucleic acids by cohesive end ligation. This process can be successfully carried out with a collection of nucleic acids comprising a complete spectrum of various sequences at the ends to be ligated, thereby providing a universal means for such ligation.

## BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the use of a poly U primer to synthesize double-stranded molecules with promoters.

Figure 2 shows the use of a 1st strand poly U primer to allow linear isothermal amplification prior to transcription.

Figure 3 shows the use of a 2nd strand ribonucleotide primer to allow linear isothermal amplification prior to transcription.

Figure 4 shows the division of nucleic acid constructs into two groups with non-complementary ends.

Figure 5 shows division using three base permutational tails.

Figure 6 shows that the oxidation of RNA prevents primer-independent cDNA synthesis by Reverse Transcriptase.

Figure 7 shows that the addition of ribonucleotides to the 3' ends of oligonucleotides inhibits subsequent dG-tailing by Terminal Deoxynucleotidyl Transferase.

Figure 8 shows the relationship between the efficiency of dG-tailing and the number of rNTPs at the 3' end.

Figure 9 shows the effect of primers with ribonucleotides on RNA yields.

Figure 10 shows the effect of 2' analogues on dG-tailing.

Figure 11 shows the effect of primers with 2' analogues on RNA yields.

Figure 12 shows the effect of primers with 2' analogues on RNA yields using multiple targets.

Figure 13 shows the effect of primers with 2' analogues on RNA yields using multiple targets after a process that did not involve Terminal Deoxyribonucleotidyl Transferase addition.

## BACKGROUND OF THE INVENTION

This is a continuation in part of U.S. Application No. 09/896,897 that disclosed numerous methods for the introduction of UDTs (Universal Detection Targets) and UDEs (Universal Detection Elements) into a library of nucleic acids. One of the intrinsic properties of nucleic acids is the potential for: a) functioning as a template for the generation of a complementary copy; and b) functioning as a primer to initiate such copying activity. In the course of carrying out the amplification of nucleic acids, one function may be desirable and the other function may be undesirable for a particular nucleic acid. For instance, the well known primer-dimer problem in PCR is the result of an unextended primer being inappropriately used as a template. There are other systems where this potential duality of functionality may also be problematic.

For instance, as described in the application cited above, amplification of a library of eukaryotic mRNA is often carried out by priming with oligo T primers that initiate synthesis by binding to the 3' poly A tails of mature mRNA. In this way, the mRNA that is present as a 3-5% minority in total RNA can be preferentially used as a source of templates without purification of the mRNA from the bulk RNA. However, there may be inappropriate priming by other nucleic acids present in the reaction mixture. The non-mRNA portion that makes up the majority of total RNA is composed of two sources: a) discrete nucleic acids such as snRNA, t-RNA and rRNA species; and b) heterologous RNA from unprocessed RNA and excised intron segments.

Even though these RNAs are not likely to be used as templates for oligo T directed synthesis due to their lack of poly A, they all possess 3' ends that can potentially act as primers. In a well known example of this, the lifecycle of a retrovirus is initiated by a priming event using the 3' end of a host cell tRNA to generate a viral cDNA copy. As such, the discrete sequences at the 3' ends of tRNAs, snRNAs and rRNAs may be able act as primers at random sites in the

multiplicity of nucleic acid sequences that make up total RNA. This potential is further increased by the use of the low stringency conditions that are necessitated by the low Tm of A:T base pairing between the oligo T primer and poly A tail. Additionally, the huge variety of sequences at the 3' ends of heterologous RNA present in a total RNA preparation could be viewed as a collection of random primers even though they are larger than the hexamers, octamers and decamers traditionally used as random primers. Lastly, the 3' ends of mRNA itself can act as primers; these ends can be discrete poly A sequences from intact mRNA or they can be random 3' end sequences of RNA fragments derived from scissions in mRNA during isolation processes.

Thus, even when it is desired to generate cDNA copies by priming with exogenously added oligo T, endogenous RNA primers present in the total RNA preparation may also participate in priming events. This effect can potentially lead to at least two negative consequences. Firstly, when eukaryotic mRNA is used as a template by these inadvertent primers, priming events can take place at sites other than the 3' poly A end, thereby creating incomplete copies. Secondly, the presence of primers that can act at random implicitly negates the specificity for selective copying of mature mRNA that was conveyed by the use of poly A; instead, the total RNA population is now a potential source of templates for cDNA synthesis. This population can be carried through later steps of labeling and/or amplification and create a higher potential for noise and background when hybridized to a target or array of targets.

In addition to the nucleic acids present in a sample to be amplified, the exogenous primers that are an intrinsic part of most amplification methodologies can contribute to inadvertent side reactions. Reference has already been made to the primer-dimer reaction of PCR, but even methods that primarily depend upon generation of multiple copies through RNA transcription can have problems from an excess of primers used in preliminary steps. For instance, Baugh et al., (2001, Nucl. Acids Res 29; e29 ) reported that when using an oligo T primer with

a T7 RNA polymerase for amplification of an RNA library using the method described by Van Gelder et al. (1990, Proc. Natl. Acad. Sci. USA 87; 1663-1667), substantial levels of RNA synthesis would be produced even in the complete absence of target input. They speculated that even though the reaction mixture had been purified from its initial components, there was sufficient carryover of unused primers to allow this *de novo* synthesis. Indeed, when the initial concentration of primers was reduced in this system, there was a substantial drop in target-independent synthesis. This side reaction is probably mediated by the previously observed ability of phage RNA polymerases to utilize 3' single-stranded ends as sites for initiation of RNA synthesis (Schenborn et al., 1985 Nucl. Acids Res. 13; 6223-6236, Krupp, 1989 Nucl. Acids Res. 17; 3023-3036). The presence of this synthesis most likely occurs even when there are target molecules present in the reaction mixture, thus leading to potential competition with the target dependent synthesis and also production of nucleic acids that may participate in further reactions downstream thereby contributing to non-specific background material.

Although a reduction in primer concentration showed a beneficial result, there are considerations that limit the application of this solution. In the presence of large amounts of target sequences, an excess of primers has to be maintained to insure that a copy is likely to be made of each of the initial targets. This can become very important for RNA species that are present in only low amounts within a library population. Secondly, when lower amounts of targets are used, kinetic effects come into play and larger amounts of primers are required to efficiently drive the reaction.

## DETAILED DESCRIPTION OF THE INVENTION

In the present invention, novel compositions and methods are disclosed for controlling the extendability of various components used in copying or amplification steps, thereby providing an enhancement of the specificity of the nucleic acids produced by these processes as well as potentially increasing the amount of such products.

One aspect of the present invention discloses a method where nucleic acid targets are treated with a reagent such that the nucleic acids are capable of functioning as templates but their 3' ends are incapable of being extended. These treatments can be carried out by a variety of chemical and enzymatic means. Examples of enzymatic means for RNA can include but not be limited to the poly A polymerase addition of ribonucleotide analogues such as cordycepin triphosphate or 3' aminoadenosine, or the T4 RNA ligase addition of a polynucleotide with a non-functional 3' end. Examples of enzymatic means for DNA can include the Terminal Deoxynucleotidyl Transferase addition of a terminator molecule such as dideoxy or acyclonucleotide terminator or the T4 DNA ligase (for double-stranded DNA) or T4 RNA ligase (for single-stranded DNA) mediated addition of an oligonucleotide with a non-functional 3' end. A template dependent DNA polymerase can also be used to add a terminator molecule such as dideoxy or acyclonucleotide terminator to the 3' ends of RNA or DNA by providing oligo T or permutational sequence oligomers with blocked ends that can hybridize near the 3' ends of the nucleic acid and serve as templates for incorporation of a terminator by the DNA polymerase. An example of a chemical means for inactivation can include the periodate oxidation of RNA. These treatments should eliminate the ability of the 3' ends of these nucleic acids to be extended while at the same time preserving their capabilities as templates.

Combinations of enzymatic and chemical methodologies may also be used. For example, after synthesis of a cDNA copy, terminal transferase may be

used to add a ribonucleotide to the 3' end followed by a periodate treatment. The methods of the present invention may also be combined with other procedures. For instance, U.S. Application No. 09/896,897 has previously disclosed addition of non-inherent UDTs to target nucleic acids to provide a primer binding site. The present invention discloses that prior to making copies of a target nucleic acid, procedures can be applied to prevent the use of a target molecule as a primer. Thus, in an example of U.S. Application No. 09/896,897, where a population of bacterial RNA molecules or fragmented eukaryotic mRNA was used as a substrate for addition of a UDT, the present invention can be applied by modifying the 3' end of the added non-inherent UDT to eliminate priming capability. For example, this may be carried out following the addition of a UDT, by subsequently adding a terminator nucleotide analogue by poly A polymerase. On the other hand, it may also be carried out in conjunction with the UDT addition step itself. For instance, a non-inherent UDT may be ligated to the 3' end of a target molecule where the UDT is a nucleic acid oligonucleotide with a 5' end that is functional for ligation and a 3' end that has an amine group instead of a free hydroxyl. In another exemplary method, addition of a UDT by Terminal Deoxynucleotidyl Transferase (TdT) or Poly A Polymerase may take place with a mixture of normal nucleotides and terminator nucleotides. The adjustment of the ratio of these components can determine the particular average length that is achieved before a terminator is likely to be added on.

The carryover of primers can be problematic in other systems as well. For example, the addition of non-inherent UDTs to the ends of nucleic acids as primer binding sites for amplification purposes in U.S. Application No. 09/896,897 has been cited above. Related methods have also been disclosed in US Patent No. 6,197,554 and Patent Application No. WO 00/75356. After synthesis of a first copy, two forms of primers will be present in the reaction mixture: extended primers derived from the use of target molecules as templates and excess unextended primers that remain in the reaction mixture. Although separation of the extended and unextended primers can be carried out by various physical

means, there may still be some carryover of unextended primers. However, since only the extended primers are desired to act as targets for the addition of a non-inherent UDT in later steps, addition of a UDT sequence to unextended primers by Terminal Deoxynucleotide Transferase or ligation can be deleterious since these products can ultimately form non-target derived double-stranded transcription units. When the primers used for first strand cDNA synthesis comprise an RNA polymerase promoter, addition of a UDT to an unextended primer can mediate conversion of a structure (Promoter-OligoT) into (Promoter-OligoT-UDT). After binding and extension of a second primer that is complementary to the UDT, a double-stranded transcription unit can be formed that synthesizes OligoT-UDT transcripts. On the other hand, if Oligo T is used to prime first strand cDNA synthesis, addition of a UDT to an unextended primer mediates conversion of (OligoT) into (Oligo T-UDT); second strand synthesis mediated by binding and extension of a second primer that contains a promoter sequence can lead to formation of a double-stranded transcription unit with the structure (Oligo T-UDT-Promoter) that synthesizes UDT-OligoA transcripts. In either case, such minimal transcription units should be extremely efficient in generating RNA products. As such, these non-target derived transcription units can act as competitive inhibitors by sequestering and/or using up transcription resources that would otherwise be available for the target-derived transcription units, thereby decreasing the net yield of amplification of target-derived sequences. In addition, if further amplification steps are carried out with this material, they will be virtually indistinguishable from target-derived products except that their small size should again lend itself to more efficient synthesis than the larger target-derived transcripts. And lastly, the presence of non-target derived labeled material can only add noise and increase the background when the library is used as a probe pool.

As such, another aspect of the present invention discloses novel methods and compositions that allow template dependent extension of primers to take place while limiting the ability of template independent extension to take place

with primers that have not participated in a template dependent extension event. One means by which this can be carried out is by the use of novel primers that comprise one or more nucleotides at their 3' ends that are other than deoxyribonucleotides.

For example, systems have been cited previously where amplification can be carried out after addition of a non-inherent UDT by means of Terminal Deoxynucleotidyl Transferase. One of the characteristics of this template-independent polymerase is that the nature of the nucleic acid of the template affects its ability to be used as a substrate for extension. Thus for instance, although the enzyme can efficiently add a ribonucleotide to the end of a DNA oligonucleotide, further additions are done poorly if at all. (J. Biol. Chem 276 ; 31388-31393.) In this case, a ribonucleotide triphosphate can serve as a donor substrate for incorporation, but a ribonucleotide at the terminus of an oligonucleotide is not efficiently recognized as an acceptor substrate. At the same time it should be noted that the presence of ribonucleotides at the 3' end of a primer is not a bar to extension by most template dependent DNA polymerases since RNA priming is a major mechanism for initation of DNA strand synthesis events *in vivo*. As such, novel primers with ribonucleotide termini will be able to be extended when synthesizing cDNA copies from RNA templates. At the conclusion of the reaction there is a chemical difference between extended and unextended novel primers due to the terminal ribonucleotide of a novel primer being converted into an internal ribonucleotide. As such, when Terminal Deoxynucleotidyl Transferase is used in a reaction mixture, there will be a differential ability to add to extended novel primers that now comprise a deoxyribonucleotide at their termini compared to unextended novel primers that still comprise a ribonucleotide moiety.

When the extended novel primers are used as templates in later steps, it is preferred that a polymerase be used that has sufficient reverse transcriptase activity that it can traverse the segment of the extended novel primer comprising

internal ribonucleotides, or that a mixture of polymerase is used where at least one DNA polymerase has reverse transcriptase activity. Any of the above methods may be used alone or they may be used in combination. For instance, a reaction mixture with novel primers comprising at least one RNA moiety at their 3' ends can undergo a degradative treatment (such as periodate oxidation) after a cDNA synthesis step and prior to terminal addition to insure that even low levels of Terminal Deoxynucleotidyl Transferase addition to unextended novel primers do not occur.

In addition to the use of a ribonucleotide, it is a subject of the present invention that one or more analogues of a deoxyribonucleotide can be used at the 3' end of a novel primer if: a) the analogue does not provide an efficient substrate for extension by at least one template independent polymerase; and b) the analogue can be used effectively for extension by at least one template dependent DNA polymerase. Since the characteristic that distinguishes ribonucleotides from deoxyribonucleotides is the absence of an OH group at the 2' position of the deoxyribose, nucleotide analogues that have substitutions for the OH group at this site were tested for the inability to be extended by a template-independent DNA polymerase, Terminal Deoxynucleotidyl Transferase. Indeed, as will be shown in a later Example, we have found that similar to ribonucleotide substitutions, primers with 2' fluoro and 2' O-methyl analogues at their 3' ends substantially lack the ability to be extended by Terminal Deoxynucleotidyl Transferase, while still being capable of extension by a template dependent DNA polymerase.

After synthesis of a cDNA copy, removal of the initial RNA template strand by alkali treatment can be carried out if desired, with the understanding that this treatment will also cause breakage in any ribonucleotide moieties in a chimeric primer used for cDNA synthesis. If it is desired to retain the chimeric primer sequences, an alternate treatment such as RNAse H digestion may be used. On the other hand, if 2' fluoro and 2' O-methyl analogues were used instead of

ribonucleotides in the primers, alkali may be used without breakage since these analogues are not alkali labile.

As disclosed in U.S. Application No. 09/896,897, ligases can be used as reagents to add UDTs to nucleic acids and can be considered to be template independent polymerases since ligases can add multiple nucleotides to a nucleic acid target. The present invention also finds utility in distinguishing between extended and unextended primers in the course of an addition reaction mediated by such ligations. For instance, a primer that comprises a terminal ribonucleotide changes its chemical properties as well as its enzyme substrate properties by being extended. In this instance, a terminal ribonucleotide would be susceptible to the periodate reaction discussed previously. On the other hand, once such a primer becomes extended, the terminal ribonucleotide is transformed into an internal ribonucleotide which is not susceptible to the periodate reaction. Thus, after extension with novel primers comprising a ribonucleotide at a 3' end, treatment with periodate should inactivate ligatability for only the unextended primers as the extended primers would remain unaffected; carrying out a subsequent ligation reaction should then selectively add the non-inherent UDT to extended primers.

The novel primers of the present invention may be homogeneous in nature or they may be of a heterogeneous or chimeric nature. For instance, methods have been described above for eliminating addition of non-inherent UDTs to unextended novel primers by ligase or terminal transferase through the use of a ribonucleotide or a 2' nucleotide analogue at the 3' end of a novel primer. Thus, if desired, the primer can be chimeric and comprise normal deoxyribonucleotides with one or more ribonucleotides or nucleotide analogues at the 3 ' terminus and still allow the selective process outlined above to be carried out. However, the nature of the nucleotides that are not at the 3' end may be as desired by the user. As such, a primer used for 1$^{st}$ strand synthesis can be homogeneous and be comprised entirely of ribonucleotides. For example, a

cDNA strand can be synthesized from poly A mRNA targets by using an Oligo ribo U primer; 2nd strand synthesis can be subsequently carried out with a deoxyribonucleotide primer comprising a promoter sequence. An illustration of this process is given in Figure 1. Removal of of the original RNA template can be carried out by RnaseH (step C of Figure 1) although the sequences opposite the ribo U primer will remain since this segment comprises double-stranded RNA. As described above, procedures can be carried out such that addition of a UDT by either ligation or terminal addition to unextended 1st strand primers should be substantially reduced or inhibited while allowing addition to extended copies (step D of figure 1). When the extended primer comprising ribonucleotides is subsequently used as a template to synthesize a complementary strand, polymerases with reverse transcriptase activity will be required if the RNA primer segment of the extended primer is desired to be copied. For instance, if a 2nd round of amplification is desired, preservation of the UDT sequence could be advantageous. Chimeric primers with ribonucleotide substitutions in the 5' region of a promoter sequence for first or second strand synthesis, may also be used. For instance, McGinness and Joyce have described substitutions of ribonucleotides into a number of positions in T7 promoter constructs that have little or no effect on transcription (2002 J. Biol. Chem. 277; 2987-2991

The methods disclosed for the prevention or inhibition of addition of UDTs to unextended primers may find particular utility with primers that are attached to beads or other solid matrices. For kinetic reasons, there is usually a vast excess of primers compared to the available targets that are eventually bound and extended. In single phase systems, the bulk of these unextended primers can be removed by various means after completion of strand extension steps. However, as described above, even comparatively small amounts of leftover material carried through into later steps can be deleterious. In contrast to single phase systems, extended and unextended primers bound to solid supports are never separated from each other and further synthetic steps take place in the presence of both unextended primers and extended primers. Consequently,

effects from unextended primers on solid supports can have even more drastic consequences than those seen in single phase systems where the majority of the unextended primers had been removed. As such, the methods described above find a special utility in dealing with the large number of unextended primers on solid supports that remain in the reaction mixtures, by inhibiting or eliminating their ability to participate in further reaction steps.

For example, latex and magnetic beads with oligo T segments are commercially available from a variety of sources. Application of the present invention can be carried out by addition of one or more ribonucleotides or nucleotide analogues to the 3' ends of such beads, thereby transforming the primers bound to the beads into chimeric primers of the present invention. For instance, a commercial lot of Oligo T beads can be modified by Terminal Deoxyribonucleotidyl Transferase mediated addition of ribonucleotides or nucleotide analogues prior to using the beads for cDNA synthesis. On the other hand, primers on supports can be made *de novo* by using appropriate phosphoramidites to synthesize oligonucleotides with appropriate groups at their 3' ends using any particular desirable sequence. These primers can subsequently be attached to beads or some other solid support. The beads with modified primers can then be used in any of the ways disclosed above where only the extended primers should participate in further synthetic steps.

Additionally, when polymeric sequences are being used as primer binding sites, a common modification of this system is the use of so-called "anchored primers". In this variation, the primers comprise a set of primers with permutational identities of the base or bases at the 3' ends. Thus for example, if poly A sequences are the desired targets, a set of 24-mers with the formula 5'-$T_{23}N_1$-3' where $N_1$ is G, A or C can be used as primers. Thus rather than being a single primer sequence, a primer set is used that comprises a mixture of three oligomers: 5'-$T_{23}$ G-3', 5'-$T_{23}$A -3' and 5'-$T_{23}$C-3'. More than one position may also be used for anchoring. For example, a set of 24-mer primers could be used

with the formula 5'-$T_{23}N_1$ $N_2$-3' where $N_1$ is G, A or C and $N_2$ is T, G, A or C. This formula would describe a set of 12 different primer sequences (3 x 4 permutations). By using sets of anchored primers, priming events can preferentially take place at the junction where a poly A tail had been appended to the discrete nucleotide sequence of a eucaryotic mRNA. This is in contrast to a completely homopolymeric oligo T where priming events could take place at any segment of the poly A tail. The present invention may make use of this system by using nucleotides or nucleotide analogues with the above described properties as the terminal permutational nucleotides. Thus for example, in a set with the formula 5'-$T_{23}N_1$-3', $N_1$ could be rG, rA or rC. In another example, a set with the formula 5'-$T_{23}N_1$ $N_2$-3' could have the permutational base identities described above, but the particular nucleotides used for $N_1$ and $N_2$ are 2' Fluoro analogues.

It should also be noted that the use of nucleotide analogues at the 3' ends of primers has also demonstrated an unexpected property. The system described by Eberwine et al. (op. cit.), uses the original RNA templates as a source of *in situ* primers for 2$^{nd}$ strand synthesis instead of addition of UDTs for primer binding sites. However, even though this method does not entail an extension step after 1$^{st}$ strand synthesis, we have surprisingly found that replacement of a normal primer with a primer comprising nucleotide analogues as the 3' end eventually results in more synthesis of RNA product than the same series of reactions with the normal primer. Although not desiring to be bound to any particular theory, this beneficial effect may be a result of increased stability at the 3' end that is endowed by the presence of the analogues. It is therefore a subject of the present invention that primers with modifications at a 3' end may be used in various promoter driven systems to increase the net amount of synthesis when amplifying a library of nucleic acids.

Nucleotide analogues at the 3' ends of primers have been disclosed recently (US Patent Application Nos. 2003/0044817 and 2003/0077609) to

increase discrimination between target sequences and non-target sequences. However, in the present invention, for library amplification where common sequences of a library of multiple species of nucleic acids are the targets for primer binding, this discrimination is not as relevant and the dual properties of: a) discrimination between extended and unextended primers; and b) increased synthesis of a library, are the beneficial effects that are sought after and achieved by the use of the nucleotide analogues at the 3' end of primers.

Another aspect of the present invention discloses novel methods for amplification of a library. In U.S. Patent Application No. 09/896,897, U.S. Patent No. 6,197,554 and Patent Application No. WO 00/75356, all of which are hereby incorporated by reference, RNA transcription from double-stranded DNA templates was used to provide high levels of amplification of a library of nucleic acids. These were all step-wise procedures where a single cycle of amplification was carried out. Further degrees of amplification may be achieved by repeating one or more of the steps of these procedures during a single cycle or by using the RNA products to start a new cycle. The present invention discloses that when a library of nucleic acids that have a UDT at each end is produced, the amplification can proceed through a continuous isothermal mode as well as the previously described stepwise methodology. By these means higher levels of amplification can be achieved with fewer manipulative steps.

Thus, for example, a conserved consensus sequence (an inherent UDT) may be used as a primer binding site for primers comprising promoter sequences. Consensus sequences have been previously used to provide global Polymerase Chain Reaction (PCR) amplification of bacterial DNA to generate an unlabeled collection of various PCR products (U.S. Patent No. 5,523,217, hereby incorporated by reference). These products were not used as hybridization probes. Instead, the particular pattern of a variety of PCR products was sufficient to generate a "fingerprint" that would be characteristic of the particular bacterial DNA starting material.

In contrast to this prior art, the present invention discloses an isothermal mode of amplification using RNA promoters and the generation of labeled products. Labeling of the nucleic acid products can be carried out by the inclusion of either labeled NTP or dNTP precursors. Alternatively, RNA or DNA products may be isolated after completion of a reaction and a separate cDNA or transcription labeling reaction carried out. This labeled material can then be used for hybridization to arrays as described previously to quantify sequence representation.

Also, as noted in U.S. Application No. 09/896,897, non-inherent UDTs may be added to a library of target nucleic acids, either prior to, or after a copying reaction. For example, when a first nucleic acid strand is synthesized from poly A mRNA and a non-inherent UDT is added, a system can be set up at this point where all the elements needed for $1^{st}$ strand synthesis, $2^{nd}$ strand synthesis and RNA transcription are supplied together and isothermal amplification of the library can occur. By using primers for $1^{st}$ strand synthesis that comprise RNA promoter sequences, anti-sense RNA products can continuously be formed. Contrariwise, by using primers for $2^{nd}$ strand synthesis that comprise RNA promoters, RNA products that reflect the orientation of the original mRNA population can continuously be formed.

Continuous or isothermal processes that use some of the elements described in these processes, i.e. transitions between DNA and RNA species, have been previously described in U.S. Patent No. 5,409,818, U.S. Patent No. 5,399,491 and Patent Application No. WO 92/08800, all of which are herein incorporated by reference. These three procedures (referred to as NASBA, TMA and 3SR respectively) share the same general principles, although there are some minor differences in the particulars of how they are carried out. However, all of these procedures also share a characteristic in that they are used only for amplifying specific sequences or a multiplex of a small number of desired specific

targets. In contrast, in the present invention the incorporation of a non-inherent UDT allows a generalized non-specific amplification of nucleic acids comprising the UDT sequences regardless of their particular native sequences. For example, if a library of DNA has a UDT ligated to each end of a series of fragments, the entire library can be amplified. On the other hand, the example above with Poly A mRNA and a non-inherent UDT can allow a selective amplification of a generalized population derived from mRNA. It should also be noted that the primers with nucleotide analogues at the 3' ends that have been disclosed above may also find application in a continuous mode of amplification of a library of nucleic acids.

Recently, novel methods of RNA promoter driven amplification have been described where only dNTPs (Haydock, U.S. Patent No. 6,531,300) or a mixture of dNTP and one or two species of NTPs (Haydock and U'ren, U.S. Patent Application No. 2003/0050444) are provided, both of which are hereby incorporated by reference. These methods are based upon the principle that although by definition, RNA polymerase synthesizes RNA transcripts, it is known that under special circumstances DNA polynucleotides can be synthesized from RNA promoters by an RNA polymerase. As described in the above references, these circumstances can be chemical in nature by replacing Mg++ with high levels of Mn++ or they may also be of genetic nature by mutating the RNA polymerase at appropriate sites. In any of the various embodiments of the present invention, RNA promoters and polymerases may be used for synthesis of DNA or chimeric DNA/RNA molecules using the methods described above or by any other method that provides this property.

In addition, U.S. Patent No. 6,531,300 and U.S. Patent Application No. 2003/0050444 exploit these circumstances to provide an isothermal mode of amplification for selected specific nucleic acid sequences. Thus, it is an object of the present invention that rather than amplifying specific target sequences, the above cited methods may be used for isothermal amplification of a library of

various nucleic acid sequences by using either inherent or non-inherent UDTs as primer binding sites to synthesize DNA or chimeric DNA/RNA molecules by means of RNA polymerase.

The system described in U.S. Patent Application No. 2003/0050444 (termed Logarithmic Isothermal DNA Amplification or LIDA) makes use of primers that have three segments in the following order:

5' Promoter-Homopolymer-Discrete 3'

The first segment (Promoter) comprises the sequence transcription of chimeric DNA/RNA by an RNA polymerase. The second segment (Homopolymer) is designed to insure that when the RNA polymerase initiates a transcription reaction, the immediate template segment will consist of DNA nucleotides. The third segment (Discrete) directs selectivity of binding of the primer to discrete target sequences.

One of the examples that has been disclosed in U.S. Patent Application No. 09/896,897 is the use of poly A RNA (an inherent UDT) to synthesize cDNA copies which were further processed by Terminal Deoxynucleotidyl Transferase mediated addition of a homopolymeric sequence (a non-inherent UDT) thereby generating a library of nucleic acids that have homopolymeric UDTs at each end. Thus, the steps used to carry out the LIDA process may be used in the present invention with primers that have only two segments and lack the discrete nucleic acid sequence required by the LIDA disclosure, the primers thereby having the following structure:

5' Promoter-Homopolymer- 3'

The homopolymeric binding sequence may be the same on the 5' end of one strand and the 5' end of the second strand, and a single primer used for isothermal amplification. Alternatively, a different homopolymeric binding sequence can be present at the 5' end of one strand versus the 5' end of the second strand, with more than one primer carrying out the amplification. The homopolymeric sequences of the primers may be complementary to an inherent

UDT (such as a poly A segment) at one end of the amplicons and complementry to a non-inherent (homopolymeric) UDT added on to the other end of the amplicons or the primers may be complementary to homopolymeric UDTs that have been incorporated into each end of the amplicons.

Other means of continuous isothermal amplification have been described by Engelhardt et al., in European Application number 0 667393 A2 (published August 16,1995, incorporated herein by reference). The present invention discloses that the addition of one or more UDTs to nucleic acids could allow the amplification of a wide variety of nucleic acid sequences by these methods. For instance, in one embodiment of Engelhardt et al., the use of primers comprising RNA segments is disclosed to allow a series of binding and ext4ension events to take place using a single template molecule. In this particular method, treatment with Rnase H allows regeneration of a primer binding site. This method has also been subsequently described in U.S. Patent No. 5,824,517 and U.S. Patent No. 6,251,639, both of which are incorporated herein by reference.

Figure 2 shows a series of events that could take place using this method. It should be pointed out that the initial product used in for the isothermal amplification of Figure 2 is the final product of step F shown in Figure 1. As described previously, the inclusion of a polymerase with reverse transcriptase and strand displacement activity could utilize the RNA segment of the first primer as a template (step A of Figure 2) thereby providing an RNA/DNA hybrid segment. Removal of all or part of the poly U RNA segment by Rnase H (step B) would generate a binding site for another poly U RNA primer. Binding and extension of this RNA primer (step C) could result in a in displacement of the original $1^{st}$ cDNA strand. The displaced cDNA strand could be converted into double-stranded form (Structure D2) if there was $2^{nd}$ strand "GGGGGGGGG-Pro" primers still present in the reaction mixture. On the other hand, the other product (structure D1) of step D would be identical to the product of step A of figure 2 and as such, further rounds of Rnase H digestion, primer binding and strand

extension can result in a series of displacemnt and accumulation of more copies of structure D2.

This method provides various means for amplification. If a source of second primers is omitted during the series of digestions, binding and extension reactions, there will be a continuous generation of single-stranded product. This in itself could be labeled if desired and hybridized against an array. On the other hand, if an RNA promoter sequence is also included in the primer for 2$^{nd}$ strand synthesis (as depicted in figures 1 and 2) there is also the potential that after conversion to a double-stranded form, a transcription reaction can also be carried out to provide further levels of amplification. It should also be noted that although an RNA primer was used for 1$^{st}$ strand synthesis in Figure 2, the opposite arrangement could have been used if desired and the 1$^{st}$ strand primer could have comprising deoxyribonucleotides coding for an RNA promoter and an oligo T segment and the 2$^{nd}$ strand could have comprised a riboG segment. This could have been combined with previous aspects of the present invention and the 1$^{st}$ strand primer could be a chimeric primer comprising one or more 2' nucleotide analogues at the 3' end of the primer. This arrangement and a potential series of events are shown in Figure 3. A succession of further binding and extension events can take place with ribo G primers binding to the single-stranded oligo C UDT of the product of step F of figure 3 with chimeric oligo T primers converting displaced single strands into double-stranded forms as described previously.

Also, although the example shown in Figure 2 used a homogeneous ribonucleotide primer, this unity is not an essential feature and the primer could have been chimeric and comprised deoxyribonucleotides as well. In essence, the minimal necessity for the system described above would be a segment that is sufficiently long enough that Rnase H digestion would allow another primer /extension event to take place. It is also a further consideration of the present invention that amplification could be carried out where the primers for both 1$^{st}$

and 2nd strand synthesis comprise RNA segments that are substrates for Rnase H action.

Another aspect of the present invention discloses the universal addition of double stranded segments by means of cohesive end ligation. In previous art, the universal ligation to ends of target molecules with undefined sequences takes place either by T4 RNA ligase, which adds to single-stranded targets, or by the use of T4 DNA ligase in a blunt end ligation reaction. Both of these processes are slow and require long incubation times (overnight is recommended in some cases) to insure adequate completion of the reactions. On the other hand, cohesive end, or as it is sometimes called, "sticky end" ligation, can be remarkably fast. For instance, commercially available kits are provided by a number of manufacturers that require only a five minute incubation period for ligation. However, in prior art, these kits have been used only for the purpose of joining discrete cohesive ends to complementary ends and as such ligation takes place only between fragments that have defined complementary sequences at their ends. Thus, these kits have been mainly applied to the joining of restriction enzyme fragments for gene construction. Another application of these kits can be for the addition of a UDT to fragments with variable terminal sequences after addition of an artificial discrete end. For instance, Rabbani et al. (in U.S. Application No. 09/896,897), Schmidt and Mueller (1996 Nucleic Acids Research 24; 1789-1791) and Komura and Riggs (1998 Nucleic Acids Research 26; 1807-1811) have disclosed the addition of sequences to a library for cohesive end ligation by means of terminal deoxynucleotidyl transferase or template-independent additions carried out by some DNA polymerases.

In contrast to this art, the present invention discloses novel nucleic acid constructs and novel methods of their use that can add desirable sequences by means of cohesive end ligation to a variety of single-stranded nucleic acid sequences without a necessity for artificially adding sequences by polymerases. The present invention discloses the use of constructs that comprise a

combination of selected conserved sequences and permutational single-stranded ends that can provide for universal ligation to a collection of nucleic acid targets of unknown or varied sequences. These Universal Ligation Elements (ULEs) may comprise any desirable functionality. For instance, the ULEs can comprise Universal Detection Elements (UDEs) that provide promoter sequences or primer binding sequences for the synthesis of copies of the nucleic acid targets. Other examples of useful functionalities that can be conveyed by ligation with a ULE can include but not be limited to restriction enzyme sites, capture sequences, capture moieties and signal moieties. The method of the present invention is based on permutational principles similar to those used with anchored Oligo T primers. In that method, a discrete sequence (oligo T) has permutations in one or two nucleotides at the 3' ends to provide universal initiation of cDNA synthesis at or near the junctions of poly A sequences in eukaryotic mRNA. As a result of the permutational design, a primer will always be provided in the reaction that is perfectly complementary to the junction, regardless of the identity of the first or second discrete nucleotide after the poly A sequence. In the present invention, binding of a set of nucleic acid constructs with permutational overhangs to a collection of single-stranded nucleic acid targets with various sequences will allow linkage of a discrete sequence to the ends of these targets regardless of the particular sequence at the termini of these targets.

Although it is possible to have a single base overhang that could be functional in the present invention, the digestion of nucleic acids with various restriction enzymes has demonstrated that single base extensions can exhibit very low efficiencies of re-ligation. For instance, catalogues of vendors of restriction enzymes such as New England Biolabs ( Beverly, MA) show efficiencies that can be as low as less than 5% . On the other hand, restriction enzymes that leave two or more nucleotide overhangs are typically very efficient (>95%) in re-ligating properly. As such, in the present invention, it is preferred that the overhang comprises at least two bases. On the other hand, permutational variations reduce the particular concentration of each individual

primer and increase the number of individual species that have to be made. Thus with a single base variation, there are 4 potential sequences and a single species represents 25% of the population. As the number of permutational positions rises this number changes accordingly: with two nucleotides, there are 16 variations, with three nucleotides there are 64 variations and with four nucleotides, there are 256 variations. Thus although higher numbers may be used, it is preferred that two to four permutational positions be used for the constructs of the present invention.

When carrying out ligation of the constructs to the single-stranded nucleic acid targets, there is also the possibility of the constructs ligating to each other rather than to the targets. In such cases, it may be desirable to remove self-ligated constructs from the remaining part of the population by using any of a variety of methods that have been used in the art for separating unincorporated primers from various extension or amplification processes. On the other hand, it may be advantageous to introduce steps to ensure that the addition of the ULEs occurs only with the target nucleic acids. For instance, there may be incomplete removal of self-ligated ULEs or there may be incorporation of more than one ULE into the terminus of a target nucleic acid. These might be events that would be detrimental to steps that occur later. As such, it may be desirable to alter the ends of ULEs so that they are unable to participate in blunt end ligations with each other. Prevention of cohesive end ligations between ULEs could be carried out by dividing the sixteen possible species of a set of ULEs with two-base, single-stranded permutational tails into two groups wherein there is no complementarity within the members of a single group. An example of a possible arrangement for this is shown in Figure 4 where six of the permutations are included in Group A and six of the permutations are included in Group B. It should be noted that the last four of the possible permutations can not be dealt with in this way, and are not included in these groups, i.e. the permutations ending in GC, CG, AT and TA are intrinsically self-complementary. To ensure that targets with these sequences at their 3' ends are still substrates for ligation

of a ULE, a potential solution is to use three-base permutations for this particular group. This is illustrated in Figure 5 with Groups C and Group D. A combination of Group A, Group B, Group C and Group D effectively comprises a collection of ULEs that can be ligated to all potential 3' end sequences. Two ways that these groups could be used are for: (a) target pool division; and (b) sequential ligations. For target pool division, a target pool can be divided into two portions. A first portion of the target pool can be ligated to a mixture of Group A and Group C and a second portion of the target pool can be simultaneously ligated to a mixture of Group B and Group D. Once the ligation steps are complete, the un-ligated ULEs are removed and the ligated products are re-pooled together. This is the simplest method, although a maximum of 50% of the targets will be ligated to the ULEs. For many purposes, this may be adequate. Conversely, a more complete ligation may be obtained by performing sequential ligations. First, the targets are ligated with a mixture of Group A and Group C. Subsequently, un-ligated ULEs are removed, and a mixture of Group B and Group D is added along with the appropriate reagents for carrying out a second ligation reaction. In this way, all of the termini of the targets are exposed at one point or another to a ULE with a complementary overhang.

A similar system may be designed that provides a set of nucleic acids that may be ligated to 5' ends. For example, an mRNA population from a bacterial source can simultaneously or sequentially go through the addition of a non-inherent UDT to the 5' end and the addition of a non-inherent UDT to the 3' end. These UDT at each end may comprise identical sequences, complementary sequences or they may comprise different sequences that may be used to distinguish between 5' and 3' ends.

The presence of a defined conserved sequence at each end of a collection of single-stranded nucleic acids allows a variety of further synthetic steps to be carried out. For example, a mRNA population comprising an inherent UDT, i.e. the poly A tail, can be used for synthesis of a collection of cDNA copies.

The cDNA population may then be used as substrates for the addition of a non-inherent UDT by the use of the present invention. 2nd strand synthesis can then be carried out by extension of a 3' end of a ULE after ligation. If preferred, the strand of the ULE that remains unligated may be removed and a separate primer added for 2nd strand synthesis. If a phage promoter is included in either a primer used for 1st strand or 2nd strand synthesis, the collection of nucleic acids can be used to carry out a transcription reaction.

In another example, RNA from a bacteria population is used as a substrate for 3' addition of a first non-inherent UDT using the present invention. Primer extension is then carried out using one strand of the ligated nucleic acid construct to form a cDNA strand. The RNA templates are removed and the same procedure is carried out to add a non-inherent UDT is added to the 3' ends of the cDNA targets. An amplification reaction can then be carried out, using primers that are complementary to the first and second non-inherent UDT's. Amplification systems that might be useful in the present invention can include but not be limited to thermocycling systems such as PCR or any of a variety of isothermal systems. Examples of the latter can include but not be limited to 3SR, TMA, NASBA, SDA, LIDA and Hairpin Mediated Amplification. SDA has been described by Walker et al., (1992 Proc. Nat. Acad. Sci. USA 89; 392-396) and Hairpin mediated amplification has been disclosed in US Patent Application No. 09/104067 filed on June 24, 1998 both of which are incorporated by reference.

In another example, RNA from a bacterial population is used as a substrate for addition of a non-inherent UDT by the methods of the present invention. The population is then hybridized to an array of complementary sequences on low density or high density microarray chips. Quantifying the amount nucleic acidscaptured on each site of the array is then carried out by hybridizing a labeled nucleic acid complementary to the non-inherent UDT.

The source of nucleic acids to be used in the methods of the present invention may comprise RNA or DNA that has been isolated from a biological source. Such sources can include phage particles, viral particles, eukaryotic cells and bacterial cells. On the other hand, nucleic acids that are derived from a one or more in vitro copying or amplification steps may also be sources of material. Thus for example, after isolation of mRNA from a biological source, cDNA copies can be synthesized and then used in accordance of the practices of the present invention. In another example, mRNA from a biological source is amplified according to the method of Eberwine to provide a collection of RNA products that have oligo T at their 5' ends and a diverse variety of sequences at their 3' ends. The methods of the present invention can be applied for addition of a primer binding site to the 3' ends by using the methods described previously for Univeral Ligation Elements (ULEs). The library can then be utilized in another round of linear amplification using a primer with a promoter for either first strand or second strand synthesis.

In most cases, the nucleic acids that are amplified for use in microarray analysis are labeled with a signal generating moiety (for example by Cy 3) or by a ligand that can be used to bind a signal generating moiety (biotin for example). However, after a series of reactions is carried out to provide this labeled material, a common preliminary step is to quantify the amount of labeled probe. In some cases, a user may find that the net yield of these process has resulted in an insufficient amount of probe to carry out the experiment. This may because there was only a marginally sufficient amount of starting material in the first place or there were losses at one or more points during the procedure. At this point, the user is usually faced with a choice of whether to proceed with a suboptimal amount of labeled probe or to go back to the original source. In the latter case, this may not even be practical if there was a limited amount of material in the first place.

In another embodiment of the present invention it is disclosed that labeled material intended for use in an array does not have to be used as a probe but can instead be used as a source of template material for carrying out further amplification procedures. Thus for instance, if the methods of the present invention have been used to generate a labeled library by a promoter in the second strand synthesis to produce sense RNA, the same primer that was used for the original first strand synthesis could be used with the labeled material to synthesize 1$^{st}$ strand cDNA and carry out a further cycle of amplification. On the other hand if a promoter was used in the 1$^{st}$ strand, thereby generating labeled antisense RNA, the same primer that was used for the second strand synthesis could be used to initiate cDNA synthesis from the labeled template. When a transcription is carried out in this new cycle, inclusion of labeled nucleotides could generate a sufficient amount of labeled probe to carry out the intended experiment with a microarray without having to go back to the original source.

## EXAMPLES

### Example 1

### Modification of the 3' ends of an RNA library.

On the day of use, a 200mM stock solution of sodium periodate (Sigma-Aldrich, St: Louis, MO) was prepared in nuclease-free water (Ambion, Austin, TX). To prepare a working solution, 10µl of 200mM sodium periodate was added to 990 µl of 10mM sodium acetate. 100 µg of total RNA from human liver (Stratagene, La Jolla, CA) in nuclease-free water was mixed with an equal volume of 2mM sodium periodate in 10mM sodium acetate, incubated for 60' on ice and protected from light. The RNA was purified using an RNeasy column (Qiagen,Valencia, CA) according to the manufacturer's protocol. The integrity of the modified RNA was verified by agarose gel electrophoresis and staining with SYBR Gold (Molecular Probes, Eugene, OR). First strand cDNA was synthesized from 10µg of modified and un-modified human liver RNA templates using Superscript II RNaseH$^-$ reverse transcriptase (Invitrogen, Carlsbad, CA) according to the manufacturer's specifications, in the absence or presence of $dT_{24}$ primer. RNA templates were removed by alkaline hydrolysis in 0.6M NaOH, 4mM EDTA (2x) at 37°C for 15' followed by neutralization with 6M acetic acid (20x). cDNA was then purified on MinElute columns (Qiagen,Valencia, CA) and eluted in 10µl buffer EB following the manufacturer's directions. DNA electrophoresis was carried out in a 1% agarose 0.5 xTBE gel, stained with SYBR Gold and viewed with a Digital Image Station (Kodak, Rochester, NY).

**Results**

The results of this experiment are shown in Figure 3. Lanes 2 and 4 demonstrate the ability of both treated (lane2) and untreated (lane 4) RNA to act as templates for cDNA synthesis with oligo T primers. On the other hand, although there is extensive cDNA synthesis in the absence of exogenous primers

for the untreated RNA (lane3), there is almost a complete absence of primer independent synthesis for the periodate-treated RNA (lane 1).

**Example 2**

**Blockage of tailing after addition of ribonucleotides.**

Tailing reactions (5µl) were carried out with 100 pmoles of $dT_{21}$ oligonucleotide primers in 1xTdT Buffer (Enzo Life Sciences, Farmingdale, NY) and 1mM cobalt chloride for 60' at 37°C, in the absence or presence of 1mM UTP and 15U Terminal Deoxynucleotidyl Transferase (TdT). The source of TdT was a recombinant bovine clone developed at Enzo Life Sciences (Farmingdale, NY) using standard methods to define the number of units of TdT activity. At the end of the tailing reaction, volumes were then increased to 50µl with the addition of 1XTdT buffer and cobalt chloride in the absence or presence of 1mM dGTP and 15U TdT. After a 15' incubation at 37°C, the reactions were quenched with 1µl of 0.5M EDTA and placed on ice. Aliquots (5µl) of each reaction were mixed with 5µl of Gel Loading Buffer II (Ambion, Austin, TX), denatured by heating for 10' at 70°C and nucleic acid strands separated by 15% PAGE containing 7.5 M Urea. A $dT_{24}$ primer (lane 7) and a 38 bp mixed sequence primer (lane 8) are included as size markers. Nucleic acids were visualized by staining with SYBR Gold and quantified with the Digital Image Station.

**Results**

The results of this experiment are shown in Figure 4. Essentially lanes 2 and 4 are the same and show that the untreated oligonucleotides (lanes 1 and 3) are substrates for the incorporation of numerous dG nucleotides after TdT treatment. It can also be seen that there is a spectrum of products encompassing different size lengths after dG addition. In contrast, lane 5 shows the effect of UTP addition by TdT. Although there is a size shift compared to lanes 1 and 3, the effect is modest and shows that only a few ribonucleotides were added to the oligonucleotides in this reaction. More importantly, in lane 6 it can be seen that

there is essentially no further increase in size when dG is attempted to be added after the ribonucleotide modification.

**Example 3**

**Comparison between one and three ribonucleotides for inhibition of dG addition.**

Since the product of the TdT mediated addition of U to the primers in the example above probably represents a collection of primers with different numbers of ribonucleotides present instead of a single discrete species, primers were synthesized with either 1 or 3 ribonucleotides already in place at the 3' end. Phosphoramidites for inclusion of ribonucleotide moieties were obtained from Glen Research (Sterling, VA). The sequences of these primers are as follows, where the 5' end comprises a T7 RNA polymerase promoter sequence and the 3' end is complementary to a poly A segment:

**PRO - $T_{24}$ =**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG TTT TTT TTT TTT TTT TTT TTT TTT-3'

**PRO - $T_{23}U_1$ =**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG TTT TTT TTT TTT TTT TTT TTT TTU-3'

**PRO - $T_{21}U_3$ =**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG TTT TTT TTT TTT TTT TTT TTT UUU-3'

Tailing reactions (8µl) were carried out with 80 pmoles of either PRO-$T_{24}$, PRO-$T_{23}U_1$ or PRO-$T_{21}U_3$ in 1x TdT buffer, 1mM cobalt chloride and 24U TdT for 15' at 37°C, in the presence or absence of 0.05mM dGTP. Aliquots of each reaction

were denatured as described in Example 2 and the nucleic acids were separated by 10% PAGE containing 7.5 M Urea. Nucleic acids were visualized by staining with ethidium bromide and quantified with the Digital Image Station.

**Results**

The results of this experiment are shown in Figure 5, where the gel results are shown in the top part. As described above, this gel was then digitally quantified and the averaged results are shown in the lower image. As a control, lanes 1, 3, 5, 7 and 9 show the size positions of the oligonucleotides in the absence of dGTP . As expected, the normal primer (PRO-$T_{24}$) shows extensive addition of dG by TdT (lane 2 compared to lane 1). On the other hand, the chimeric oligonucleotides (carried out in duplicate) are not as efficient substrates. For instance, it can be seen in lanes 4 and 6 that even with a single ribonucleotide at the 3' end (PRO-$T_{23}U_1$) there is extensive blockage of terminal addition, and the majority of the oligonucleotides remain in the position of the untreated oligonucleotide. In lanes 8 and 10, this effect can be seen to be intensified by the use of multiple ribonucleotide substitutions (PRO-$T_{23}U_1$), where it can be seen that there is little if any addition by TdT.

**Example 4**

**Chimeric primers with ribonucleotides at the 3' end show increased transcription.**

**(a) 1st strand synthesis**

A recombinant clone (ATCC 87482) for poly $A^+$ bacterial Lys A was obtained from the American Tissue Culture Collection (Manassas, VA) and used to prepare mRNA *in vitro* using standard procedures. 1µg of purified poly$A^+$ Lys A RNA (1,272 nt) was mixed with 50pmole of the PRO-$T_{24}$ primer or the PRO-$T_{21}U_3$ chimeric primer (used in Example 3) in a 10µl solution, heated for 10' at 70°C, followed by the addition of 9µl of a RT premix containing 2 µl of 10X Stratascript RT Buffer (Stratagene, LaJolla, CA), 1µl of 10mM dNTP mix, 2µl of 100mM DTT,

1µl of 30U/µl RNase Inhibitor (Eppendorf, Boulder, CO) and 3µl of nuclease-free water. After a 3' preincubation at 42°C, 1µl of 200U/µl StrataScript RNaseH$^-$ reverse transcriptase (Stratagene, LaJolla, CA) was added and the reactions were incubated at 42°C for 60' to synthesize first-strand cDNA. The final concentrations of components in the 20µl reaction mixtures were; 50mM Tris (pH 8.3), 75mM KCl, 3mM $MgCl_2$, 10mM DTT, 0.5mM of each dNTP, 1.5 U/µl RNase Inhibitor, 2.5□M primer and 50ng/µl RNA. After heat inactivation of the RT (70° C for 10'), 1µl of 5U/µl RNaseH (New England Biolabs, Beverly, MA) was added and samples incubated for 20' at 37°C. First strand cDNA was purified by MinElute columns using the manufacturer's protocols. The eluted first-strand cDNA (20µl) was dG-tailed for 15' at 37°C in a 40µl reaction containing 30U TdT, 0.5mM dGTP , 1X TdT buffer and 1mM cobalt chloride. The 3'-ends of the tailed first-strand cDNA were blocked by the terminal addition of dideoxy-ATP with TdT. In this case, 40µl of termination mix containing 1 xTdT buffer, 20mM $MgCl_2$, 5mM dideoxy-ATP (Amersham Biosciences, Piscataway, NJ) and 15 U TdT was added and the samples incubated for 15' at 37°C. The single-stranded, G-tailed and terminated cDNA was purified by MinElute columns and eluted with 20µl EB.

**(b) 2$^{nd}$ strand synthesis**

For second strand (SS) synthesis, either water (2µl) or 25 pmoles of $dC_{12}$ (2µl) was added to 10µl of the eluted cDNA, incubated at 70° C for 10', cooled to 42° C for 2', followed by the addition of 8µl of SS mix containing 4 µl of 5X SS Buffer (125mM Tris (pH 7.5), 500mM KCl, 25mM $MgCl_2$, 25mM $(NH_4)_2SO_4$ and 5mM DTT), 2µl of 10mM dNTP mix, 1 µl of nuclease-free water and 1µl of 5U/µl Klenow (New England Biolabs, Beverly, MA). After 60' at 42° C, the double-stranded cDNA was purified by MinElute columns. The final concentration of components in the 20µl SS reactions were; 25 mM Tris (pH 7.5), 100mM KCl, 5mM $MgCl_2$, 5mM $(NH_4)_2SO_4$, 1mM DTT, 1mM of each of the dNTPs, 5U Klenow, cDNA and 0 or 2.5□M primer.

**(c) Transcription**

One quarter of the eluted double-stranded cDNA (5µl) was subjected to an *in vitro* transcription (IVT) reaction for 16 h at 37°C using the BioArray High-Yield RNA Transcript Labeling Kit (Enzo Life Sciences, Farmingdale, NY) with unlabeled nucleotides only (3.75mM of each NTP). RNA was purified with RNeasy columns (Qiagen, Valencia, CA) according to the manufacturer's directions and eluted with 40µl of nuclease-free water. One-tenth (4µl) of each of the eluted RNAs was mixed with an equal volume of Gel Loading Buffer II, denatured at 65°C for 5' and electrophoresed on a 1.2% agarose, 0.5 xTBE gel. After staining with SYBR Gold, the RNA was visualized with the Digital Image Station and quantified using Kodak software. Since the target in this example was a single species, only the stained material in the size position for the LysA was used for digital quantification.

**Results**

Duplicate sets of the above reactions are shown in Figure 6. As in figure 5, the gel image is shown with the digital quantification below it. Although both the standard primer and the chimeric primer reactions resulted in generation of a substantial amount of transcription products in the absence of second strand primer addition, it is clear in this example that there were substantially more transcription products for the chimeric primer reactions compared to the standard primer reactions.

**Example 5**

**Blockage of addition of dG by chimeric oligos with 3'ends containing either 2'-Fluoro or 2'-O-Methyl nucleotide analogues.**

Phosphoramidites for inclusion of 2' Fluoro and 2' O-methyl U analogues and universal support were obtained from Glen Research (Sterling, VA). The sequences of these primers are as follows, where the 5' end comprises a T7

RNA polymerase promoter sequence and the 3' end is complementary to a poly A segment:

**PRO - $T_{21}OMe_3$**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG TTT TTT TTT TTT TTT TTT TTT XXX-3'

where X = 2'-O-Me-U

**PRO - $T_{21}Fl_3$**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG TTT TTT TTT TTT TTT TTT TTT YYY-3'

where Y = 2'-Fl-U

**PRO - $T_{22}Fl_2$**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG TTT TTT TTT TTT TTT TTT TTT TYY-3'

Where Y = 2'-Fl-U

Tailing reactions (8µl) were carried out with 50 pmoles of PRO-$T_{24}$ (from Example 3), PRO - $T_{22}Fl_2$, PRO - $T_{21}Fl_3$ or PRO - $T_{21}OMe_3$ in 8µl reaction mixtures containing 1x TdT Buffer, 1 mM cobalt chloride, 0.05 mM dGTP, 24U TdT, for 15' at 37°C. As controls, reactions were carried out as above with the addition of EDTA to 31mM. Aliquots of each reaction were denatured as described previously and nucleic acids were separated by 10% PAGE containing 7.5 M Urea. Nucleic acids were visualized by staining with ethidium bromide.

**Results**

The results of this experiment are shown in Figure 7. The reactions with EDTA serve as negative controls to show the positions of the oligonucleotides in the absence of terminal addition. In the positive control with a standard primer, there can be seen a substantial shift in position after dG addition (lane 1 compared to

lane 2). However, with either 2 (lane 3) or 3 (lane 5) substitutions with 2'fluoro analogues, there is a substantial drop in the efficiency of dG addition with the 3 substitutions (PRO - $T_{21}Fl_3$) being slightly more effective for blockage than the 2 substitutions (PRO - $T_{22}Fl_2$). In this example, another 2' analogue was also tested. Lane 7 shows that the 2' O-methyl substitution also showed inhibition of TdT tailing.

**Example 6**

**Transcription products from extended 2'-Fluoro or 2'-O-Methyl chimeric primers.**

**(a) 1st strand synthesis**

20 pmoles of PRO-$T_{24}$ (from Example 3), PRO - $T_{21}Fl_3$ (from Example 5) or PRO - $T_{21}OMe_3$ (from Example 5) primers were annealed to 0.5 µg poly $A^+$ LysA (described in Example 4) in a 10µl reaction volume. First-strand cDNA synthesis was then carried out as described previously in Example 4, except that the 3' preincubation step was omitted and the RT was included in a10µl RT Mix . Removal of RNA was achieved by base hydrolysis, (0.3 mM NaOH and 2 mM EDTA for 20' at 37°C) followed by neutralization with acetic acid. After purification through MinElute columns, the first-strand cDNA was tailed and terminated as described in Example 4., except that dCTP (0.5mM) was substituted for dGTP and only 2.5 units of TdT was used.

**(b) 2nd strand synthesis**

After purification, the dC-tailed, terminated first strand cDNA (18 µl) was annealed with 20 pmoles of $dG_{12}$ primer (2µl) and second-strand cDNA was synthesized as described in Example 4, except that the final reaction volume was 30µl instead of 20µl.

**(c) Transcription**

Transcription and analysis of duplicate reactions were carried out as described in Example 4.

**Results**

The results of this experiment are depicted in Figure 8 with the digital quantification shown below the gel results. Since this was a single species, again only the stained material in the region for the LysA transcript was used for quantification. Also, since the experiments were performed in duplicate, the numerical results are the result of averaging between the duplicates. It can be seen that the presence of the 2' Fluoro analogues in the primers resulted in an increase in the amount of transcription compared to the standard primer reactions. It can also be seen that the overall yield of the 2' O-methyl is lower. The presence of transcripts in both the 2' Fluoro and the 2' O-methyl reactions demonstrates that a) the 3' ends of the analogues can be extended and b) the segment of the 1$^{st}$ cDNA strand containing these analogues can be used as templates when carrying out 2$^{nd}$ strand synthesis.

**Example 7**

**Use of ENZO system with 2'-Fluoro chimeric primers for transcription of 4 different poly A targets.**

20 pmoles of either the T7 promoter-$dT_{24}$, or the T7 promoter-$dT_{21}$2'-F-$dU_3$ primer were annealed to 10 ng of four target poly-$A^+$ RNA transcripts of 750, 1000, 1400 and 2000 nts, (10 ng each) in 10µl volumes. First strand cDNA synthesis was carried out as described in Example 6. RNA was removed as described in Figure 6, and after purification through MinElute columns the first strand cDNA was tailed with dCTP rather than dGTP. After termination with ddATP and purification, the dC-tailed, terminated first strand cDNA (18 µl) was annealed with 20 pmoles of $dG_{12}$ primer (2µl) and second strand cDNA synthesis

was carried out as described in Example 6, except that complete Exo$^-$ Pol I was used instead of Klenow. One quarter of the double-stranded cDNA was subjected to an in vitro transcription (IVT) reaction at 37°C for 4 h using the BioArray High-Yield RNA Transcript Labeling Kit (Enzo) with unlabeled nucleotides (3.75mM of each NTP). RNA was purified by the RNeasy Kit according to the manufacturer's instructions. One-tenth of each of the eluted RNAs was denatured at 65°C in a formamide loading buffer, electrophoresed on a 1.2% agarose gel (0.5 xTBE), stained with ethidium bromide, visualized with a Kodak Image Station and quantified using Kodak software. Each individual band was measured separately.

**Results**

These reactions were carried out in duplicate and the transcription results are depicted in Figure 9 with the quantification shown below the gel picture. It can be seen that the results shown previously with the LysA is not an artifact of using that particular RNA as a substrate since all 4 of the test RNAs in this Example were able to show enhanced production of transcription products when the 2' Fluoro primers were used.

**Example 8**

**Use of modified Eberwine system with 2'-Fluoro chimeric primers for transcription of 4 different poly A RNA targets.**

20 pmoles of either the T7 promoter-$dT_{24}$, or the T7 promoter-$dT_{21}$2'-F-$dU_3$ primer were annealed to 10 ng of four target poly-$A^+$ RNA transcripts of 750, 1000, 1400 and 2000 nts, (10 ng each) in 10µl volumes. First strand cDNA synthesis was carried out as described in Example 6. Second strand synthesis and purification of double-stranded cDNA was performed according to a protocol in the Affymetrix GeneChip Expression Analysis Manual (Affymetrix, Santa Clara, CA) for a modified version of the original Eberwine method. One quarter of the double-stranded cDNA was subjected to an in vitro transcription (IVT) reaction at

37°C for 4 hours using the BioArray High-Yield RNA Transcript Labeling Kit (Enzo) with unlabeled nucleotides (3.75mM of each NTP). RNA was purified by the RNeasy Kit as suggested by the manufacturer. One-tenth of each of the eluted RNAs was denatured at 65°C in a formamide loading buffer, electrophoresed on a 1.2% agarose gel (0.5 xTBE), stained with ethidium bromide, visualized with a Kodak Image Station and quantified using Kodak software.

**Results**

These reactions were carried out in duplicate and the transcription results are depicted in Figure 10 with the quantification shown below the gel picture. Although the effect is not as large as seen in Example 7, the results of this example show an enhancement in production of transcripts of all 4 RNA products even though the reaction never involved a Terminal Deoxynucleotidyl Transferase addition step.

**Example 9**

**Use of homopolymeric ribopolymer as a primer for 1st strand synthesis.**

In this example two primers will be used that have the following sequences:

The primer for first strand synthesis $U_{24}$ would be comprised entirely of ribonucleotides:

**$U_{24}$ =**

5' UUU UUU UUU UUU UUU UUU UUU UUU 3'

The primer for second stand synthesis would comprise an RNA promoter sequence and comprise the sequence:

**PRO – $G_{12}$ =**

5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG GGG GGG GGG G-3'

A series of reactions could be carried out as described in Example 4 where **$U_{24}$** primer was used for 1st strand synthesis and **PRO – $G_{12}$** primer would be used for 2nd strand synthesis. Also, instead of the G-tailing carried out in Example 4, dCTP would be substituted to form an oligo C 3' tail to allow a subsequent binding by the **PRO – $G_{12}$** primer to carry out 2nd strand synthesis. By having the primer in the 2nd strand primer instead of the 1st strand primer as used in Example 4, transcription would proceed from the opposite direction and RNA would be transcribed that would be in the same orientation as the original target mRNA instead of the anti-sense product made in Example 4.